ARMY RESEARCH LABORATORY

Post-Processing for the Battlescale Forecast Model and Mesoscale Model Version 5

Jeffrey E. Passner

ARL-TR-2988 **June 2003**

NOTICES

Disclaimers

Army Research Laboratory
White Sands Missile Range, NM 88002-5501

ARL-TR-2988 **June 2003**

Post-Processing for the Battlescale Forecast Model and Mesoscale Model Version 5

Jeffrey E. Passner
Computational Information and Sciences Directorate, ARL

REPORT DOCUMENTATION PAGE	*Form Approved* *OMB No. 0704-0188*

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED (From - To)
June 2003	Final	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Post-Processing for the Battlescale Forecast Model and Mesoscale Model Version 5	
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	**5d. PROJECT NUMBER**
Jeffrey E. Passner	
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
U.S. Army Research Laboratory Computational Information Sciences Directorate Battlefield Environment Division (ATTN: AMSRL-CI-EB) White Sands Missile Range, NM 88002-5501	ARL-TR-2988
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	**10. SPONSOR/MONITOR'S ACRONYM(S)**
U.S. Army Research Laboratory 2800 Powder Mill Road Adelphi, MD 20783-1145	**11. SPONSOR/MONITOR'S REPORT NUMBER(S)** ARL-TR-2988

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

The Battlescale Forecast Model (BFM) produces many forecasting parameters including temperature, pressure, dewpoint, relative humidity, wind information, as well as precipitation amounts. While these output data provide valuable weather information Tactical Decision Aids such as the Integrated Weather Effects Decision Aids (IWEDA) have a need for additional parameters such as icing and turbulence. The IWEDA generates current and forecasted impacts on approximately 70 weapon systems, such as attack helicopters and fixed wing aircraft, and personnel. Of most consequence to IWEDA, are the weather hazards on military operations. These hazards include three-dimensional (3-D) weather effects such as icing, turbulence, and clouds as well as several two-dimensional products that include surface visibility, fog, and thunderstoms. The raw output fields from the BFM and the MM5 can be used to derive these vital weather parameters.

15. SUBJECT TERMS

Mesoscale, models, BFM, MM5

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
					Jeffrey E. Passner
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER (*Include area code*)
U	U	U	U	50	505-678-3193

Standard Form 298 (Rev. 8/98)
Prescribed by ANSI Std. Z39.18

Contents

List of Figures

List of Tables

Preface

The Integrated Meteorological System, the weather system for the U.S. Army has implemented a mesoscale model and mesoscale data to be used by Air Force weather forecasters in support of Army operations. Both the Battlescale Forecast Model (BFM) and the Pennsylvania State University/National Center for Atmospheric Research mesoscale model version 5 (MM5) are used for short-term and long-term forecasts respectively. As a method to provide more detailed weather data for the Army user and the tactical decision aids, the U.S. Army Research Laboratory developed the Atmospheric Sounding Program (ASP) to assist the Staff Weather Officer. Originally, the ASP was designed to use radiosonde data to display a sounding and formulate derived weather products known as weather hazards products. In recent years, the ASP was incorporated into the BFM and MM5 data to provide forecast output of many additional weather hazards that cannot be solved numerically in the mesoscale models.

This report describes the post-processing techniques and a comparison of the BFM and MM5 products.

Executive Summary

Introduction

The U.S. Army Research Laboratory (ARL), has developed a mesoscale weather model called the Battlescale Forecast Model (BFM). The BFM provides prognostic forecast variables for a 24-h period after model initialization. However, due to Army requirements, a longer-term forecast was essential, so The Pennsylvania State University/National Center for Atmpsheric Research Mesoscale Model Version 5 gridded data are received from the U.S. Air Force Weather Agency (AFWA) to provide forecast information for up to a 48-h period. To enhance the forecasts, a post-processing package, the Atmospheric Sounding Program (ASP), has been developed to run with data from both models. The ASP is designed so that it can manipulate gridded data from either model and provide detailed forecast information of weather hazards such as icing, turbulence, clouds, surface visibility, fog, and thunderstorm probability.

Purpose

This report describes the basic meteorological theory applied by the ASP for post-processing and the different weather hazards that might interfere with military operations. The techniques used by the ASP have been designed to work with any mesoscale forecast model. The effectiveness of the BFM and MM5 weather output are analyzed and discussed in this report.

Overview

The ASP is initialized from numerical model data, such as the BFM and MM5. These data provide the forecaster and users with a detailed overview of the atmospheric conditions that might interfere with military equipment and personnel. The ASP uses these data and produces a series of weather hazards that can be used for analysis or forecasts to 72 h from the initial time of the BFM or MM5 run. Included in these weather hazards are thunderstorm probability, turbulence, icing, clouds, surface visibility and precipitation type. These meteorological parameters are later placed into a database so that other programs such as the Integrated Weather Effects Decision Aid can attain this information.

1. Introduction

The Integrated Meteorological System (IMETS) is a mobile operational automated weather data receiving, processing, and disseminating system utilized by Air Force weather forecasters in support of Army operations. The U.S. Army Research Laboratory (ARL) is supporting the forecaster to make more precise and meticulous weather decisions in the battlefield by providing weather products on IMETS. One product to assist in short-term forecasting ($\leq$24 h) is an operational mesoscale model, the Battlescale Forecast Model (BFM). For longer-term data, the Pennsylvania State University/National Center of Atmospheric Research mesoscale model Version 5 output is available from 6 to 48 h *(1,2)*.

The BFM produces many forecasting parameters including temperature, pressure, dewpoint, relative humidity, wind speed, and direction as well as precipitation amounts. While these outputs provide valuable weather information, Tactical Decision Aids (TDAs) such as the Integrated Weather Effects Decisions Aids (IWEDA) have a need for additional parameters such as icing and turbulence. The IWEDA has been developed to simplify the manner in which environmental impacts on weather systems are displayed to the user. The IWEDA generates current and forecasted impacts on approximately 70 weapon systems, such as attack helicopters, fixed-wing aircraft, and personnel. The weather hazards on military operations is of most consequence to the IWEDA and the military. These hazards include three-dimensional (3-D) weather effects such as icing, turbulence, and clouds as well as several two-dimensional products that include surface visibility, fog, and thunderstorms. The raw output fields from the BFM and the MM5 can be used to derive these vital weather parameters *(3)*.

This report is divided into the following sections, each with a different degree of detail.

Section 1 - Introduction

Section 2 – Mesoscale Models for the Army

Section 3 – Weather Hazards

Section 4 – Statistical Evaluation of Mesoscale Models and the Weather Hazards

Section 5 - Summary and Discussion

2. Mesoscale Models for the Army

Pielke describes the mesoscale as having a temporal and a horizontal scale smaller than the conventional rawinsonde network but significantly larger than individual cumulus clouds. The vertical scale extends from tens of meters to the depth of the troposphere. With a requirement to provide the Army with small-scale weather information on the order of less than 500 by 500 km, ARL implemented the Higher Order Turbulence Model for Atmospheric Circulation (HOTMAC) as their model for the IMETS platform. HOTMAC was selected since it uses Alternating Direction Implicit (ADI) numerics, which ensures numerical stability at longer time steps, because it emphasizes boundary-layer physics, and is globally relocatable and platform-independent. However, to keep the model run time as fast as possible, the model contains no cloud microphysics package or convective cloud parameterization. The model in its current configuration is only run to 24 h; however, due to planning of missions it was necessary to add the MM5 to the IMETS platform to provide forecast grids out to 48 h from the initial forecast time *(4,5)*.

2.1 The BFM

The original HOTMAC software was modified for Army use by employing a horizontal resolution of 10 km, with 16 terrain-following vertical levels and a model top 7000 m above the highest elevation on the grid. A log-linear vertical stagger is used so that there is greater resolution near the surface. The BFM has levels at 2 and 10 magl, which are the standard observing heights for temperature, humidity, wind speed, and wind direction respectively. The basic variables that are prognostically forecasted by the model are perturbation potential temperature, total water substance mixing ratio, wind speed, wind direction, pressure, soil temperature, turbulence kinetic energy and length scale, and non-convective precipitation rate (*6,7*).

As already noted, the rapid run time for the model can be attributed to a single nest and no moist physics or cumulus parameterization routines. Because of the implicit approach, time steps on the order of 200 s (at 10 km resolution) are common for typical atmospheric advective speeds and vertical motion fields. Soil temperature on five subsurface levels is solved using the heat conduction equation, while long and shortwave radiation within a single layer for a stratus cloud are calculated using the method suggeted by Sasamori. The precipitation rate is parameterized as a function of cloud liquid water using the scheme developed by Sundquist *(8,9)*.

To initialize the BFM, surface data and upper-air observations are input into the model in the area-of-interest. Additionally, the 36-h forecasted Naval Operational Global Atmospheric Prediction System (NOGAPS) package, which is issued by the Air Force Weather Agency (AFWA) via the Air Force Automated Weather Distribution System is utilized as the long-range

data that the BFM is nudged toward. The NOGAPS grid points are spaced 1° latitudinal distance apart on the mandatory pressure surfaces. Lateral and time-dependent boundary conditions (large-scale forcing) are supplied from grid-point data close to the area-of-interest taken from the NOGAPS output valid at analysis and forecast times of interest.

The BFM-generated output for the grid include the u and v horizontal wind vector components, potential temperature, and water vapor mixing ratio. These forecast fields are saved at 0, 3, 6, 9, 12, 15, 18, 21, and 24 h from the base time of the model run and placed into a Gridded Meteorological Data Base (GMDB).

In summary, the main points of the BFM are listed below:

- terrain elevation data
- graphical user interface for user input
- meteorological input data; NOGAPS, surface and upper-air data
- data analysis for model initialization and boundaries
- prognostic model run
- BFM output placed into GMDB every 3 h
- post-processing: derived products placed into GMDB every 3 h
- output and displays on map background

2.2 The MM5

The Fifth-Generation NCAR/Penn State Mesoscale Model (MM5) is the latest in a series that was developed in the early 1970s. Since then, the MM5 has evolved to the current fifth generation. The MM5 is a limited-area, non-hydrostatic, terrain-following sigma-coordinate model designed to simulate or predict mesoscale and regional-scale atmospheric circulations.

Terrestrial and isobaric meteorological data are horizontally interpolated from a latitude-longitude mesh to a variable high-resolution domain on Mercator, Lambert Conformal, or polar stereographic projection. Since the interpolation does not provide mesoscale detail, these interpolated data may be enhanced with observations from the standard network of surface and rawinsonde stations using either a Cressman or multiquadric scheme. In the MM5 there is also a program that performs the vertical interpolation from pressure levels to sigma coordinates. The sigma surfaces near the ground closely follow the terrain, while the higher-level sigma surfaces tend to approximate isobaric surfaces.

Other features of MM5 are:

- globally relocatable
- flexible and multiple nesting capability
- advanced physical parameterization
- 3-D data assimilation system via nudging
- ability to run on various platforms *(10)*

The version of the MM5 being used in this study is Version 3 from AFWA with a resolution of 15-km mesh data on 41 vertical levels. ARL receives these MM5 data in gridded binary form (GriB) for the Continental United States twice each day, which are initialized at 0600 UTC and 1800 UTC respectively. Due to computational and processing constraints, there is a 6-h stagger between the initialization valid time of the 15-km mesh and the first forecast output, thus the first MM5 forecast is a 6-h forecast. The frequency of the model output is every 3 h, for a time period of 48 h.

The current AFWA operational version of MM5 places the lowest model vertical level at 20 magl. To generate data at the standard observation heights of 10 and 2 magl, similarity theory is being used at ARL to extrapolate to these lower levels from the lowest MM5 sigma level. In this fashion, temperature, dewpoint, and wind data at levels 2 and 10 magl are produced at ARL in addition to the 41 MM5 sigma levels of data.

The parameterizations selected by AFWA with this version of the MM5 are as follows:

1. **Grell cumulus parameterization.** Designed for grid sizes of 10 to 30 km, this parameterization accounts for subgridscale convection and compensating subsidence.
2. **MRF planetary boundary-layer model.** Parameterizes the mixture of heat, moisture, and momentum in the boundary layer.
3. **Reisner mixed phase explicit moisture microphysics.** Cloud and rainwater fields and ice processes are predicted explicitly. No graupel or riming processes are calculated.
4. **Dudhia cloud radiation.** Provides solar and infrared fluxes at the ground and atmospheric tendencies resulting from the radiative processes.
5. **MM5 five-layer soil model**.

Post-processing of the MM5 at AFWA includes a number of variables and is called (MMPOST). However, given the huge size of the GriB files, and the number of parameters not needed by the Army, it was decided not to include the MMPOST variables in GriB data collection from AFWA. Additionally, many of the parameters needed by the IWEDA and other TDAs are not included in the MMPOST data, thus the Atmospheric Sounding Program (ASP), the post-processing program for the BFM is being used by the MM5 in the IMETS environment.

The main components of the MM5 are:

- terrain data
- Data ingest of surface observations, upper-air observations, and global-model data
- REGRID, interpolates the global model data to the MM5 grids
- Little_r/RAWINS, interpolated data enhanced with surface and upper-air observations
- Interpf, vertical interpolation from pressure levels to sigma coordinates
- MM5, model package run
- post-processing of MM5 data
- archive and display of data *(11)*

3. Weather Hazards

Often in weather forecasting, decisions must be made instantaneously, so it becomes beneficial to implement artificial intelligence (AI) techniques to assist in weather forecasting. The weather hazards program is not truly AI because it uses statistical data, conventional computer programming techniques, and basic meteorological calculations as a first "guess" at the hazards. As an example, the cloud forecast is based on a continuous sequence of rules that uses relative humidity data, derived lapse rates, moisture depth, wind data, time of day, seasonal influences, and location of the station. All these facts are synthesized by a set of rules to make a forecast of cloud height, ceiling height, depth of the cloud, and cloud amounts.

3.1 Turbulence

Turbulence is a state of fluid in which there are irregular velocities and apparently random fluctuations. Due to large updraft and downdraft speeds, turbulence can be expected in and near thunderstorms; thus a thunderstorm indicates that pilots must adjust their flight routes near these convective clouds *(12)*.

Forecasting clear air turbulence (CAT) is a more complicated and frustrating problem because of the small timescale and resolution that turbulence is often observed with it. Ramer correlated synoptic weather patterns to the observation of CAT. Work done by Lake in 1956, and more recently Black and Marroquin has linked calculations of kinetic energy to areas of forecasted turbulence (*13-15)*.

Theoretical studies and empirical evidence have associated CAT with Kelvin-Helmholtz instabilities. Miles and Howard indicate that the development of such instabilities require the existence of a critical Richardson number (RI) $\leq$0.25. Stull notes that the Richardson number is a simplified term or approximation of the turbulent kinetic energy equation where the RI is expressed as a ratio of the buoyancy resistance to energy available from the vertical shear *(16,17)*.

The equation is expressed below:

$$RI = \frac{\frac{g}{\theta} * \left(\frac{\partial \theta}{\partial Z}\right)}{\left(\frac{\partial V}{\partial Z}\right)^2} \tag{1}$$

where g is the gravitational acceleration, $\partial\theta/\partial Z$ is the change of potential temperature with height, and ∂V is the vector wind shear occurring over the vertical distance ∂Z.

The U.S. Navy Fleet Numerical Meteorological and Oceanography Center (FNMOC) uses the Panofsky index (PI) to forecast low-level turbulence, where the low level is considered to be below 4000 ft above ground level (AGL). The formula for this index is:

$$Panofskyindex = (windspeed\ )^2 * \left(1.0 - \frac{RI}{RI_{crit}}\right) \tag{2}$$

where RI is the Richardson number and RI_{crit} is a critical Richardson number empirically found to be 10.0 for the FNMOC data. The higher the Panofsky index the greater the intensity of turbulence at low levels *(18)*.

Investigation of pilot reports and radiosonde upper-air observation (RAOB) data showed that the Panofsky index in the lower levels and the Richardson number above 4000 ft provided the best skill scores . Meanwhile, Ellrod and Knapp 1992 listed environments where significant CAT was found to be prevalent. Their study associated vertical wind shear, deformation, and convergence into a single index. This work by Elrod and Knapp was based on the Petterssen's frontogensis equation and was ideal to utilize the gridded output of a mesoscale model. Originally, they used the nested grid model and global aviation model to develop and evaluate their turbulence index. Later, Knapp researched and validated the Turbulence Index (TI) using the 16-level BFM at ARL *(19–21)*.

Using the Panofsky index below 5000 ft AGL and the Richardson number above that level to the model top of 7000 magl, Passner found that the Panofsky index was most effective in the lowest 5000 ft while the Richardson number was generally ineffective between 5000 to 10000 ft AGL and more effective above 10000 ft AGL. The results in the Passner study indicated a need for an improved routine above 5000 ft AGL. It was determined to implement the TI above 4000 ft

AGL, since Knapp and Smith in their 1995 study were able to prove that a combination of the features of the TI and the PI provided the highest correlation coefficients *(22)*.

3.2 Icing

Icing typically occurs at temperatures between 0 and –40 °C. In the ASP, three types of icing are considered

1. rime
2. clear
3. mixed

While the four icing intensities in the ASP are

1. trace
2. light
3. moderate
4. severe icing

Since the BFM does not have a cloud microphysics package, it was determined that the best approach to the analysis/forecasting of icing was to use the RAOB icing tool developed at the AFWA in 1980. The RAOB technique uses the temperature, dew-point depression, and temperature lapse rate as a measure of instability of the layer. A study by Knapp showed that the RAOB icing tool performed with the most accuracy *(23,24)*.

The RAOB tool categorizes icing by lapse rate, temperature, and dew-point depression; the three temperature groups are: –35 to –16 °C, –16 to –8 °C, and –8 to –1 °C. These temperature classes are based on the theory of ice formation, with the first case, –35 to –16 °C, resulting in light rime icing in all cases. The middle class, –16 to -8° C, generally accounts for the mixed and rime cases, with the intensity based on the lapse rate or stability of the layer. The warmest class –1 to –8 °C, is often the temperature range when clear icing is found. A final case was added to account for severe clear icing. This situation occurs when there is a strong inversion about 100 mb above the surface so that the relatively warm water droplets spread quickly on the aircraft and cause clear icing to form.

In his study, Cornell investigated numerous soundings data and found that the mean dew-point depression for all icing types was 4.5°. Due to an underforecasting bias, this adjustment was added to the RAOB icing tool in this ARL study *(25)*.

3.3 Clouds

Numerical models often contain cloud-physics packages and cumulus-convection routines that solve for cloud heights, ceilings and cloud amounts. Since the BFM is designed to run as quickly as possible, there is currently no cloud physics package. It was decided to approach the cloud-forecasting problem with a cross between empirical techniques, statistical data, and rule-based IF-THEN sets of code.

Work done by Walcek indicated that a 2 to 3 percent increase of the relative humidity could lead to a 15 percent increase in cloud cover. His findings were employed to derive the "decision tree" or flow chart that is used to form the IF-THEN rules in the cloud program *(26)*.

As noted by Schultz, mesoscale models often have a dry bias. Schultz observed cases where relative humidity values in excess of 55 percent between 500 to 1000 mb on the Nested Grid Model were related to cloudy conditions. The BFM does not display such an extreme bias; however, clouds are often observed in layers with relative humidity well below values of saturation *(27)*.

A study using 13 runs of the BFM indicated the differences between the BFM output and observed soundings. The results are shown in table 1.

Table 1. The difference in the BFM output against the observed sounding, based on hours from initial time and different pressure levels.

Pressure Levels/Time from Initialization	**00-h Forecast**	**12-hr Forecast**	**24-h Forecast**
925 mb	3% drier	5% drier	11% drier
850 mb	3% drier	8% drier	9% drier
700 mb	2% drier	10% drier	11% drier
500 mb	3% drier	3% moister	12% drier

Table 1 illustrates what might be expected, a drier model output in comparison to measured relative humidity by the soundings. The model moisture is fairly consistent at all levels at the initial time; however the model is 5 to 10 percent drier at the 12-h and 24-h periods. These data indicate that making cloud predictions from BFM model output would require a lessening in the relative humidity values, with the greatest differences occurring at lower pressure levels and with increasing time. At the initial forecast time (00-h), 53 to 63 percent of the model runs were drier than the observed soundings. At the 12-h forecast period, the models were drier than the soundings 65 percent of the time and finally at 24 h the models were drier 71 percent of the time.

Since cumulus clouds cannot be forecasted by the BFM, convective clouds were added by an empirical method in the post-processing software. During the time of maximum heating (1100 to 2000 local), cumulus clouds were formed if the convective temperature was being approached or

exceeded. The cumulus clouds persisted only if the convective temperature was exceeded during the forecast period.

3.4 Surface Visibility

Low visibility is another example of a weather hazard that impacts military ground and air operations. In an effort to compile a database for deriving a universal visibility equation, Knapp collected 2790 observations from July 1994 to April 1995. He included station elevation, temperature and dew point, dew-point depression, relative humidity, wind speed, ceiling height, and precipitation reported as his set of variables. From the 2790 surface observations, two types of equations were formulated, which account for different conditions based on available surface observation data *(28)*.

These two equation types were:

Type 1. Ceiling known, precipitation unknown

Type 2. Ceiling and precipitation known

Screening regression techniques using stepwise procedures were used to determine the predictor values for each equation type. Once the "best" correlated predictor was found, other predictors were then included to achieve the best statistical results.

As an example, the equation developed using observations with derived ceilings with no precipitation falling:

$$\begin{aligned} VISCAT &= 7.41 + (.0005 * ELEV) - (.0088 * DEWPT) - (.0371 * RH) \\ &\quad + (.0268 * WINDSPD) + (.0044 * CIG) \end{aligned} \tag{3}$$

where VISCAT is the category of the predicted surface visibility, ELEV is the surface elevation, DEWPT is the surface dew point, RH is the relative humidity, WINDSPD is the surface wind speed, and CIG is the height of the ceiling. For each equation, empirical adjustments are made based on the ceiling and surface visibility. As an example, using eq 3, the following empirical adjustments are made:

If VISCAT=4 and CIG>=25 and RH<90, change VISCAT to 5.

If VISCAT=5 and CIG<=10 and RH>=85, change VISCAT to 4.

4. Statistical Evaluation of Mesoscale Models and the Weather Hazards

Statistical evaluation of mesoscale models is traditionally focused on the meteorological variables produced directly by the model or those solved numerically. Often these evaluations

focus on the model temperatures, winds, and heights of the certain pressure levels. While these data provide useful ways to study model performance, they often do not address the weather problems that the model users may directly face. This study will focus on the moisture fields and the weather hazards that are highly dependent on moisture output.

4.1 Evaluation Techniques for Forecasts

Two types of evaluations are done in this study.

1. "YES/NO" forecasts, where the forecast provides information if a certain weather phenomena will or will not occur. An example of this is the turbulence forecast where the user gets a simple "YES/NO" prediction if turbulence is expected or not.
2. The model output is investigated for "error" or how much the predicted value differs from the observed value.

4.1.1 Evaluation of "YES/NO" Forecasts

A contingency table (tab. 2) provides a statistical method to display answers to binary YES/NO forecasts. Some evaluation techniques include the probability of detection (POD), false alarm rate (FAR), the correct non-event (CNE), critical success index (CSI), true skill score (TSS), and bias. The calculations are based on the contingency elements listed in table 2, while the equations for the evaluation techniques are also shown.

Table 2. Contingency table for forecasted and observed weather event.

	Forecast YES	**Forecast NO**
Observed YES	A	B
Observed NO	C	D

$$POD = \frac{A}{A+B} \tag{4}$$

$$FAR = \frac{C}{C+A} \tag{5}$$

$$CNE = \frac{D}{D+C} \tag{6}$$

Donaldson developed the CSI, which considers three of the four elements in the contingency table; however, it does not take into account the D element, the null element. Hanseen and Kuipers formulated an equation that does factor in the null event, and called it the TSS *(29,30)*.

$$CSI = \frac{A}{A+B+C} \tag{7}$$

$$TSS = \frac{(AD)-(BC)}{(A+B)(C+D)} \tag{8}$$

The bias in a forecast is the ratio of the number of positive forecasts to the number of observed events as shown in eq (9)

$$Bias = \frac{A+C}{A+B} \tag{9}$$

4.1.2 Error Evaluation

The three main products used in this study to evaluate model or post-processed derived output are mean absolute difference (AD) , root-mean square error (RMSE), and correlation coefficient (CC). The equations appears below:

$$AD = \frac{\sum_{j=1}^{m}\sum_{i=1}^{n}\left|x_{o,i,j} - x_{p,i,j}\right|}{m*n} \tag{10}$$

Where

x = meteorological variable

o = observation

p =prediction of variable

i = i^{th} surface station

j = forecast day

n = number of stations,

m = total number of forecast days

Small values of AD are related to good agreements between observation and forecast.

$$RMSE = \sqrt{\frac{\sum_{j=1}^{m}\sum_{i=1}^{n}(x_{o,i,j} - x_{p,i,j})^2}{m*n}} \tag{11}$$

The values of root mean square error are proportional to those of the absolute difference. The CC is displayed in eq 12. The CC measures the strength of the relationship between two variables. When CC >0 it indicates a positive linear relationship. A value of 1.00 indicates a “perfect” correlation between the observed and predicted values of a meteorological forecast.

$$CC = \frac{\sum_{j=1}^{m}\sum_{i=1}^{n} x_{o,i,j} * x_{p,i,j}}{\sqrt{\sum_{j=1}^{m}\sum_{i=1}^{n} x^2{}_{o,i,j} * \sum_{j=1}^{m}\sum_{i=1}^{n} x^2{}_{p,i,j}}} . \tag{12}$$

4.2 Model Evaluation

The purpose of model evaluation is to investigate and study the skill of the basic model output. Knowledge of the model trends and their biases will help to understand the strengths and weaknesses of each post-processed variable such as icing and clouds.

This model evaluation in this study was conducted from October 2001 to May 2002 and involved both the BFM and MM5. There were 32 model runs completed in this study; all 1200 UTC runs for the BFM and 0600 UTC for the MM5. Surface observations were collected at random sites on the grid; however, every effort was made to pick points in different terrain regimes on different parts of the grid. The upper-air points were compared to actual upper-air observations from RAOB stations on the grid. The model runs were done in a variety of locations representing different weather conditions, with 10 of the 32 runs centered over the Chicago area to study wintertime conditions. For both the BFM and MM5, a total of 75 to 86 surface observations were studied for each forecast output time. Table 3 shows the BFM surface statistics while table 4 gives the results of the MM5 surface forecasts where dew point temperature (TD) depicts the dew point depression.

Table 3. BFM surface temperature and surface moisture errors through 24-h.

BFM Hours and Variables	Mean Absolute Difference	RMSE	Correlation Coefficient
00-h Temperature	1.93	2.55	0.96
06-h Temperature	2.22	2.91	0.94
12-h Temperature	2.08	2.71	0.95
18-h Temperature	2.47	3.33	0.88
24-h Temperature	2.86	3.58	0.90
00-h Dew point	1.04	1.74	0.98
06-h Dew point	2.08	2.71	0.94
12-h Dew point	2.82	3.80	0.91
18-h Dew point	3.06	4.12	0.80
24-h Dew point	3.55	5.30	0.81
00-h TD depression	2.06	2.91	0.75
06-h TD depression	4.06	5.12	0.59
12-h TD depression	3.76	4.83	0.73
18-h TD depression	3.38	4.31	0.59
24-h TD depression	2.86	3.87	0.55

Table 4. MM5 surface temperature, dew point, and dew-point depression statistics for model output from 06 to 36 h.

MM5 Hours and Variables	Mean absolute Difference	RMSE	Correlation Coefficient
06-h Temperature	2.34	2.98	0.85
12-h Temperature	2.42	3.07	0.95
18-h Temperature	2.68	3.48	0.93
24-h Temperature	2.31	2.98	0.94
36-h Temperature	2.25	2.76	0.97
06-h Dew point	2.42	3.31	0.94
12-h Dew point	2.02	2.74	0.96
18-h Dew point	2.35	3.10	0.94
24-h Dew point	2.33	3.13	0.94
36-h Dew point	2.33	2.95	0.93
06-h TD depression	2.11	3.16	0.66
12-h TD depression	3.13	4.01	0.81
18-h TD depression	3.37	4.42	0.78
24-h TD depression	2.52	3.46	0.72
36-h TD depression	3.02	3.97	0.86

Some of the interesting trends noted in these data are that the BFM has a much better skill score with the initial dew point field than the temperature field. This trend is not noted in the MM5, although the correlation coefficient is higher for the surface-moisture field than the temperature field. The BFM has a pronounced drop in skill by 18 h after model initialization. The MM5 output does not show a significant decline in skill in either the temperature or moisture field during the 36 h of data shown. The MM5 has its largest error at the 18-h mark in the model run, which is 0000 UTC. The BFM has its highest error and lowest skill at the 6-h time period or at 1800 UTC, possibly due to the difficulty in modeling the mixing in the lower atmosphere and all the radiation balances. For both models, the correlation coefficients are far lower for the dew-point depression than for either the temperature or dew point, most likely due to the wider variability that this parameter often has.

To show the bias of the model temperature and moisture field, a test was done to evaluate if the model was overforecasting or underforecasting the temperature and dew point. For a forecast to be overforecasting or underforecasting the temperature or dew point, the error had to be greater than +0.2 ºC between the forecast and the observation. These results are shown in table 5 where they are displayed as the percentage of the total forecasts in the full sample that the variable is overforecasted or underforecasted.

Table 5. Percentage of cases where model overforecasts or underforeacsts the surface temperature and surface dew point.

Model, Time and Variable	Overforecast (%)	Underforecast (%)
BFM 00-h Temperature	56	38
BFM 12-h Temperature	44	48
BFM 24-h Temperature	65	33
BFM 00-h Dew point	40	35
BFM 12-h Dew point	39	54
BFM 24-h Dew point	37	53
MM5 06-h Temperature	36	53
MM518-h Temperature	27	67
MM524-h Temperature	29	64
MM5 00-h Dew point	52	36
MM5 12-h Dew point	48	48
MM5 24-h Dew point	53	38

In general, the BFM appears to overforecast the surface temperature while the MM5 underforecasts the surface temperature. Conversely, the BFM underforecasts surface dew points after the 00-h forecast while the MM5 has a slight bias to overforecast the surface moisture until the 36-h forecast period.

In figure 1, the chart shows the RMSE and bias for the surface temperature from the MM5 off the site, http:/weather.afwa.mil/index.html. This plot displays temperature RMSE on the top through the 72-h forecast period of the MM5. The RMSE ranges from 3 ºC in the early forecast periods to as much as 5 ºC by the 72-h forecast. There is little significant error between the 0600 UTC and 1800 UTC forecast cycles. The lower (dashed) part of the graph shows the model bias in degrees Celsius. Generally, the MM5 underforecasts the temperature with only some occasional peaks of overforecasting displayed overnight. It is uncertain as to why these slight jumps in model temperatures occur, and these biases do not agree with the work done in the ARL study. However, the overall bias of underforecasting the temperatures is similar to the work in this report.

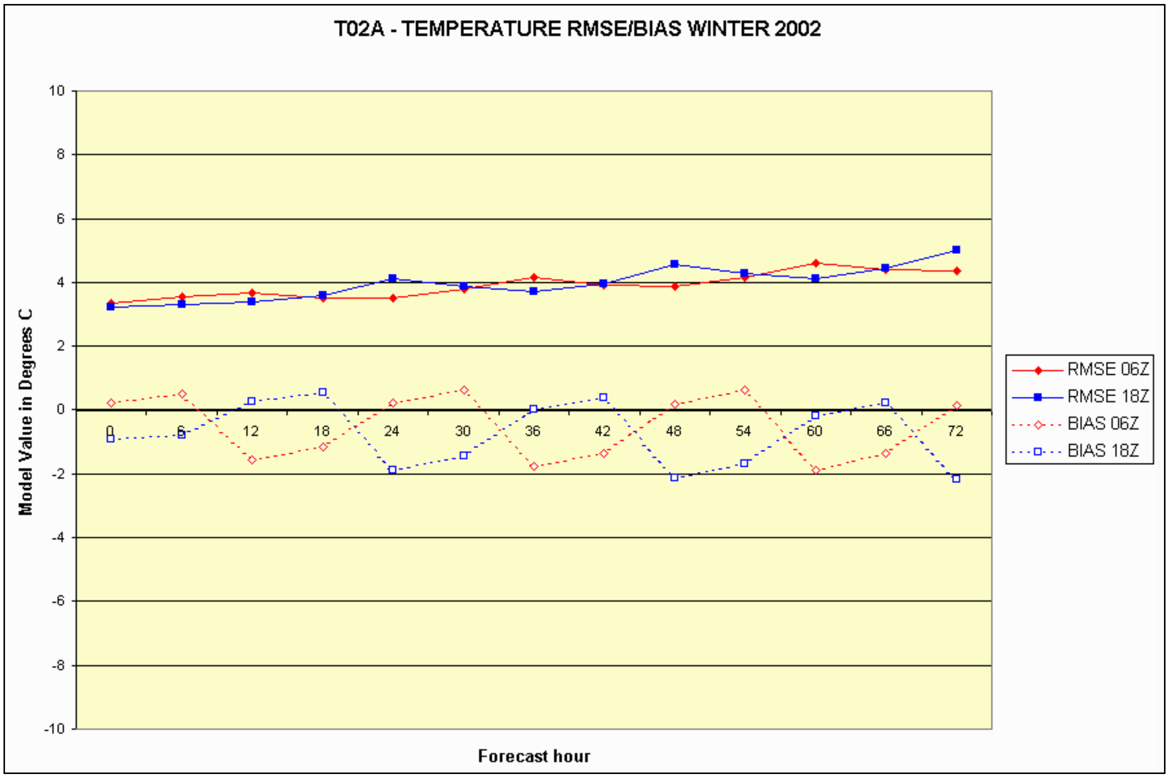

Figure 1. The temperature error in the AFWA MM5 study during the winter of 2002.

Figure 2 shows the relative humidity RMSE and bias for the MM5 at AFWA. This plot agrees with the results in table 5 since figure 2 shows the RMSE in relative humidity (percent) from the 00 to 72-h period. The errors increase from about 12 percent at the initial time to 19 percent at the end of the period. The bias is to overforecast the relative humidity through the entire forecast cycle at both 0600 UTC and 1800 UTC.

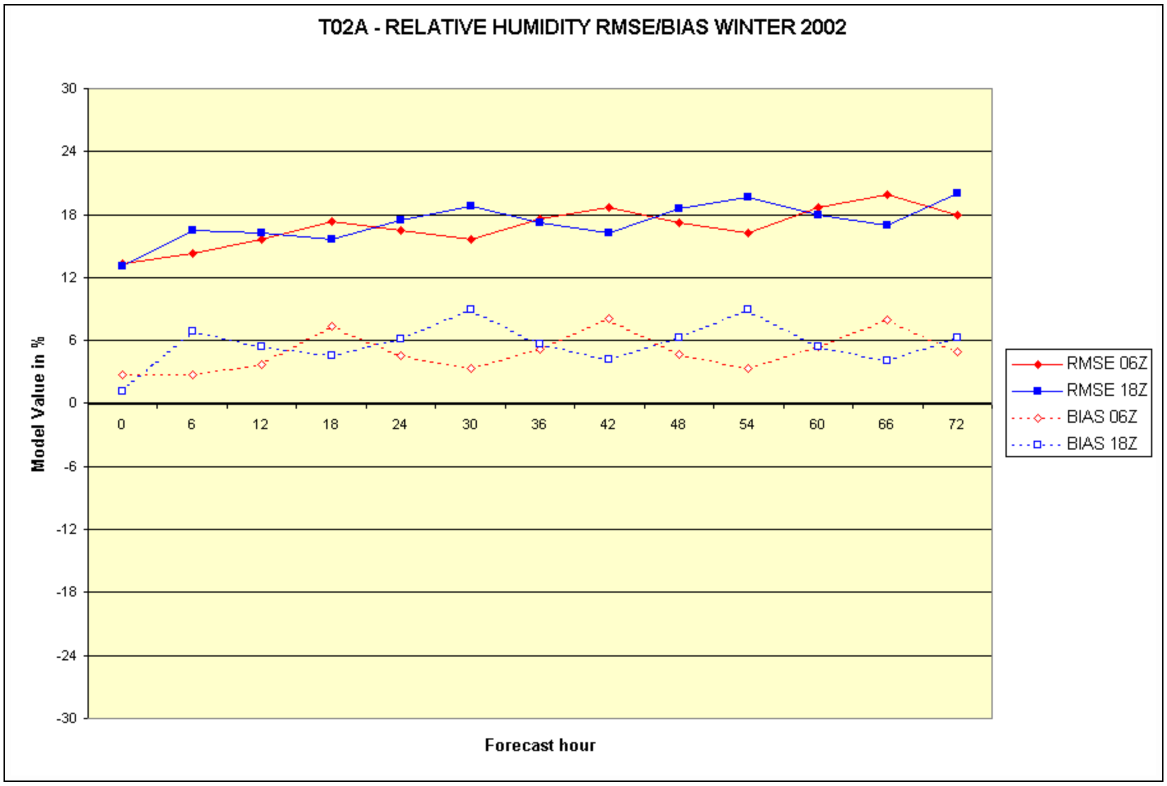

Figure 2. The surface relative humidity error from the AFWA MM5 during the winter of 2002.

Another attempt to understand model performance was to see how effective each model was in dry and moist environments. While there are no particular standards to differentiate between dry and moist surface conditions, it was determined that any dew-point depression less than 5 ºC was considered a moist surface case while any data point with a surface dew-point depression greater than 5 ºC was determined to be a dry environment. Below are the results for all model-output hours.

Dry cases

BFM temperature: Underforecasts surface temperature 70 percent of the time

MM5 temperature: Underforecasts surface temperature 86 percent of the time

BFM dew point: Overforecasts surface dew point 61 percent of the time

MM5 dew point: Overforecasts surface dew point 77 percent of the time

Moist cases

BFM temperature: Overforecasts surface temperature 65 percent of the time

MM5 temperature: Underforecasts surface temperature 60 percent of the time

BFM dew point: Underforecasts surface dew point 63 percent of the time

MM5 dew point: Underforecasts surface dew point 52 percent of the time

It is not entirely certain as to why the models underforecast the surface temperature field in the dry cases, although both models show the greatest error in this bias during the warmest hours of the day, between 1800 UTC and 0000 UTC. This indicates that in dry surface environments that a cool bias might be associated with excessive moisture at higher levels

- “mixing” problems
- a lack of a land-use model
- albedo errors
- no model knowledge of the soil moisture and soil types.

The complex interaction of model feedbacks make it difficult to understand all results such as the ones in this section; however, these results do give modelers some clues to model deficiencies or problems.

The moisture error in the dry environment appears to be greatest between 0600 to 1200 UTC. Again, it is uncertain as to why this occurs given all the factors that might be involved; however, given the time of day, the boundary layer is greatly influenced by radiational cooling and even small errors in the clouds or winds can cause large model temperature errors.

Testing of upper-air model output was also conducted in this study with comparisons between model forecasts and upper-air observations done at 1200 UTC and 0000 UTC. As displayed in table 6, the BFM temperature and moisture forecasts at 700 mb are shown where. There are 30 samples for each of the forecast hours in table 6.

Table 6. 700-mb BFM temperature and moisture results at 1200 and 0000 UTC.

BFM 700 mb	**Mean Absolute Difference**	**RMSE**	**Correlation Coefficient**
00-hr T[a] (1200 UTC)	0.90	1.22	0.89
12-hr T (0000 UTC)	1.84	2.28	0.85
24-hr T (1200 UTC)	1.94	2.24	0.94
00-hr TD[b]	3.18	4.64	0.89
12-hr TD	6.40	9.50	0.40
24-hr TD	6.43	9.04	0.64
00-hr TD-dep[c]	3.43	5.04	0.73
12-hr TD-dep	7.20	11.08	0.27
24-hr TD-dep	6.35	9.57	0.65

[a]T represents the temperature
[b]TD is the dew point
[c]TD-dep is the dew point depression

The results of the BFM output show high correlations in the temperature field but lower correlation in the moisture field. This result, the lower skill in the mid-level moisture field, is not a surprise because much of the moisture change may be attributed to moisture advection and vertical accelerations, which are not handled well by most models. Additionally, a lack of upper-air data in the initial data and interpolation errors can lead to large errors in the moisture field. Another possible problem with the mid-level moisture fields is that by the 12 and 24-h forecast period the BFM forecast has nudged strongly to the NOGAPS forecast, which at this point can be as old as 24 or 36 h. Table 7 shows the 700-mb temperature and moisture errors for the MM5 through 42 h. The number of sample is 30 for each forecast hour.

Table 7. Temperature and moisture results at 700 mb for the MM5.

MM5 700 mb	**Mean Absolute Difference**	**RMSE**	**Correlation Coefficient**
06-h T (1200 UTC)	0.88	1.11	0.98
18-h T (0000 UTC)	1.14	1.55	0.97
30-h T (1200 UTC)	1.31	1.61	0.97
42-h T (0000 UTC)	2.06	2.56	0.93
06-h TD	5.45	7.34	0.88
18-h TD	5.68	8.23	0.64
30-h TD	6.34	9.66	0.63
42-h TD	5.25	6.93	0.77
06-h TD-dep	4.32	6.20	0.78
18-h TD-dep	5.79	8.41	0.49
30-h TD-dep	6.71	8.37	0.63
42-h TD-dep	5.65	7.74	0.72

The MM5 results show the same pattern as the BFM; the temperature errors are far less than the moisture errors. The MM5 skill in the temperature forecasts decreases with time, but the RMSE and correlation coefficient are exceptional in the long-range periods of this study. However, the RMSE is high in the moisture forecast indicating that the MM5 has the same difficulties with the mid-level moisture field as the BFM.

Like the surface, it is advantageous to understand the model biases at 700 mb. Table 8 shows the bias for each of the upper-air forecast periods.

Table 8. Forecast biases for the BFM and MM5 at 700-mb.

BFM and MM5 Hours	Overforecast (%)	Underforecast (%)
BFM 00-h Temp[a]	13	70
BFM 12-h Temp	26	74
BFM 24-h Temp	22	70
BFM 00-h TD[b]	40	60
BFM 12-h TD	42	48
BFM 24-h TD	63	37
MM5 06-h Temp	31	45
MM5 18-h Temp	24	62
MM5 30-h Temp	34	66
MM5 42-h Temp	38	55
MM5 06-h TD	48	45
MM5 18-h TD	48	41
MM5 30-h TD	62	38
MM5 42-h TD	48	45

[a]Temp is the temperature
[b]TD is the dew point

Both models have a bias to underforecast the temperature at 700 mb; however the bias in the mid-level moist field is not as clear. The BFM appears to be too dry at the initial time but too moist at 24 h. The MM5 has a bias to overforecast dew point in the midlevels, although this bias is much stronger at 30-h for no known reason. As already noted, the sample size of only 30 700 mb cases may not be large enough to fully show the biases.

Figures 3 and 4 are the results of the AFWA 2002 study over CONUS using the MM5. These two charts show temperature and relative humidity errors.

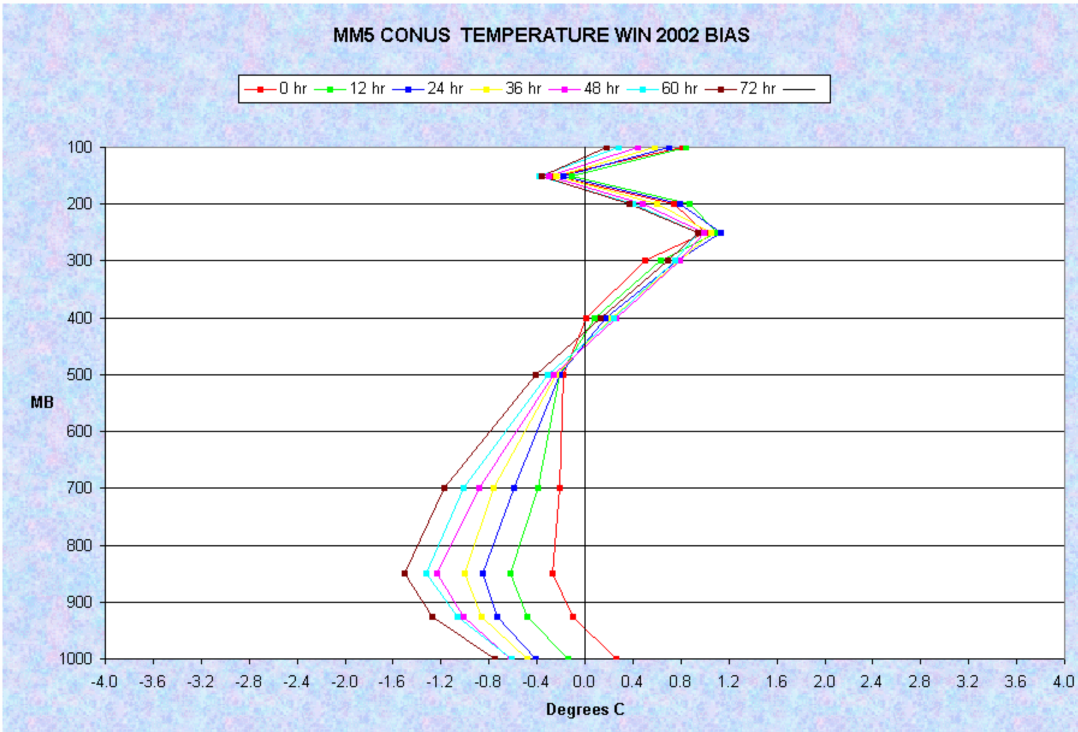

Figure 3. Temperature bias with height from AFWA 2002 study.

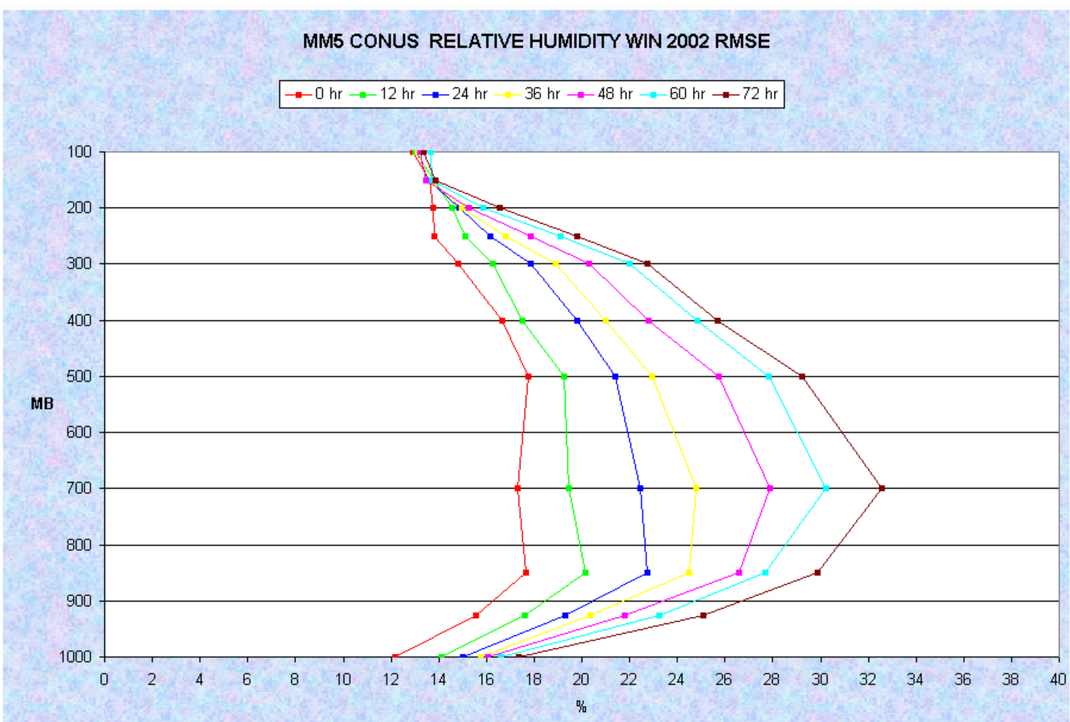

Figure 4. MM5 relative humidity errors from AFWA 2002 study.

The results in figure 3 agree with the results in table 8. The MM5 underforecasting the temperature through most of the atmosphere and overforecasting the dew points will lead to errors in the relative humidity fields at all levels as seen in figure 4. The error in the relative humidity varies from as low as 12 percent at the initial time to as much as 33 percent by 72 h.

Below are the two cases, the dry and moist cases for both models.

700-mb dry cases:

BFM temperature:	Underforecasts 700 mb temperature 76 percent of the time
MM5 temperature:	Underforecasts 700 mb temperature 62 percent of the time
BFM dew point:	Overforecasts 700 mb dew point 57 percent of the time
MM5 dew point:	no significant bias

700-mb moist cases:

BFM temperature:	Undeforecasts 700 mb temperature 64 percent of time
MM5 temperature:	Underforecasts 700 mb temperature 68 percent of the time
BFM dew point:	no significant bias
MM5 dew point:	no significant bias

These statistical studies were completed for 850, 700, 500, and 300 mb. Below are some the most important biases found during this model evaluation:

- The BFM underforecasts temperatures at all levels and hours.
- The MM5 underforecasts temperatures from the surface to 300 mb.
- The BFM has a slight bias to underforecast moisture at the surface, 850 mb and 700 mb, but overforecasts moisture at 850 mb and 700 mb at 24 h.
- The MM5 overforecasts the dew point at all levels tested and at all hours.
- In the "dry" environment, the BFM overforecasts the dew point; however, in the moist environments the BFM underforecasts the dew point from the surface to 700 mb.
- In both the dry and moist cases, the MM5 underforecasts the temperature field.
- In the dry environments, the MM5 overforecasts the moisture, however, in the moist model runs this bias is not as significant.
- Both the BFM and MM5 underforecast the surface wind speeds. This bias does not exist with increasing height.

- At the surface, the BFM underforecasts the wind speed in 62 percent of the cases, while the MM5 underforecasts the wind speed in 57 percent of the cases. Both models rarely overforecast the wind speeds.
- Both the BFM and MM5 have a lower wind-forecast skill at 0000 UTC than at 1200 UTC.
- Correlation coefficients increase with height in the wind study; however, the RMSE does not change significantly with height.

The model biases in this study are part of the model numerics, and they greatly influence the post-processing package. In the next section, a detailed study of the post-processed variables will show the relation between model numerical output and derived variables.

4.3 Weather Hazards Evaluation

Evaluation of the weather hazards is important since it gives the user a general idea of how the TDAs will perform in the battlefield. Additionally, it helps to understand how influential the model numerics are in the post-processing of the derived variables. In this section, the main emphasis will be on turbulence, icing, clouds, and visibility.

4.3.1 Turbulence Evaluation

The method used in this study to verify turbulence is to compare pilot reports (PIREPs) to model forecasts. Using the BFM and MM5 output, verification is limited to a 1-h period surrounding the model forecast time. As an example, model forecasts of turbulence at 2100 UTC are compared to PIREPs from 2030 to 2130 UTC only. Any PIREPs that included two intensities, such as light (LGT) to moderate (MDT), were classified as the more extreme intensity. As a standard, only PIREPs close in height to the model forecast were accepted. For levels below 10000 ft AGL, the forecasted turbulence had to be within 1000 ft of the PIREP. From 10000 to 20000 ft AGL, the forecast had to be within 1500 ft of the PIREP, and above 20000 ft AGL, the forecast had to be within 2000 ft of the observed turbulence.

During the winter season of 2002, model runs were made using the MM5 and BFM. All BFM runs were for 24 h, while the MM5 were used for the full 48-h forecast period. The models were run for the same area; however, a direct comparison between models was not done since the models have different initialization times and ingest different data at these times. Table 9 displays the results for the BFM and MM5 turbulence forecasts using the PI and TI combination. These results include turbulence for all levels of the atmosphere at all forecast hours for each model run.

Table 9. “YES/NO” turbulence statistics using the BFM and MM5.

Turbulence Evaluation	**BFM 2002 Study**	**MM5 2002 Study**
Samples	455	648
POD	0.66	0.75
FAR	0.27	0.26
CNE	0.59	0.53
CSI	0.53	0.59
TSS	0.25	0.28
BIAS	0.91	1.02

The results in table 9 indicate little difference in skill between the models in the POD, FAR, and CSI which would agree with the data from the wind evaluation. The results in table 10 divide the turbulence results so that the influence of the PI and TI can be examined.

Table 10. The “YES/NO” turbulence forecasts for PI and TI for all forecast hours.

	BFM	**MM5**
POD-PI	0.80	0.92
FAR-PI	0.29	0.17
POD non-event – PI	0.38	0.49
Bias- PI	1.13	1.03
POD-TI	0.60	0.73
FAR-TI	0.30	0.30
POD non-event –TI	0.61	0.56
Bias – TI	0.85	1.02

Based on the results shown in table 10, both models have a very high POD in the lowest layers. However, much of this test was done in “obvious” weather conditions; cases when turbulence was expected. There is a stronger influence in the wind speed using the PI in the lower levels, which might be expected given the squared term in eq 2. Thus, the stronger the wind, the higher the value of the PI, given the same temperature profile. Furthermore, the POD of the null event in the lower levels is only 0.38 for the BFM and 0.49 for the MM5, thus the PI does have a bias to overforecast the turbulence in the lowest 4000 ft AGL. With increasing height, the TI, which is based on convergence, deformation, and vertical motions shows slightly lower POD but compensates for that trend by having much higher skill in forecasting the non-event case.

Another interesting study involves investigating the turbulence forecasts by increasing time from the model initialization. Figure 5 shows the POD of turbulence through 48 h.

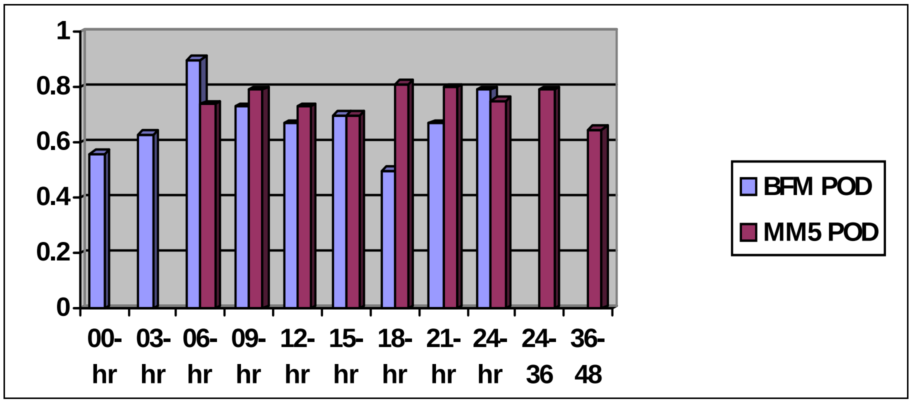

Figure 5. POD of turbulence for all levels by forecast hour in BFM and MM5.

In figure 5, there are no data for the first two time periods for the 15-km MM5 since the model does not output data before the 6-h forecast. The results do show a higher POD at the 6- and 9-hr forecast time frame. It is uncertain as to why the BFM has a low POD at the initial time, a high POD at the 6-h and a low POD at the 18-h period. The MM5 POD is steady until 36 h after model start time.

Turbulence intensity is very challenging to forecast, but the verification of this parameter is even more difficult since it depends on the pilots as noted by Kane, who reports that 72 percent of pilots send incomplete PIREPs *(31)*.

It is best to look for a broad range of error in verifying the intensity, with as many cases as possible in the correct "class." The BFM study in 2002, shown in table 11, consisted of 455 samples while the MM5, also in 2002 and displayed in table 12, had 645 samples. In the tables, the number of forecasts for each intensity is the vertical columns, while the number of observations are the horizontal rows.

Table 11. Turbulence intensity forecasts and observations for the BFM for all model levels and forecast hours.

Obs/Forecasts	None	LGT	MDT	SVR[1]	Total
None	99	44	28	0	171
LGT	49	38	21	1	109
MDT	44	53	56	1	154
SVR	3	10	8	0	21
Total	195	145	113	2	455

Table 12. MM5 turbulence intensity, all forecast hours and all levels.

Obs/Forecasts	None	LGT	MDT	SVR	Total
None	131	71	26	7	235
LGT	47	60	22	12	141
MDT	45	93	64	29	231
SVR	7	11	14	6	38
Total	230	235	126	54	645

[1]Severe

The fact that more pilots report turbulence as moderate is probably a result of pilots ignoring the no turbulence and light turbulence cases since they have minimal impact of the aircraft. It also may be a result of the large number of smaller planes involved in the study and the subjective reporting of turbulence. Additionally, the “LGT to MDT” reports are considered as moderate turbulence in this study, as a worst-case scenario. In both models, there is a bias to underforecast the turbulence intensity.

While not shown in a table or chart, another result in this study is that both the BFM and MM5 provide more accurate turbulence forecasts on days when widespread turbulence is observed, such as days with large storms and dynamical lifting. Many of the errors in the study appear to be in the cases of occasional light turbulence, which are often forecasted to be “no turbulence,” and may not have a significant error on Army aircraft. There are few forecasts of severe turbulence in the BFM, which leads to a bias of underforecasting severe turbulence events. The MM5 sample has far more severe turbulence forecasts (8 percent of sample), but it is encouraging to see that in many of the cases, the severe forecast is matched with a moderate observation.

4.3.2 Icing Evaluation

The evaluation of the model icing forecasts was completed in the same manner as the turbulence evaluation. Icing forecasts were correct if the level of the PIREPs were close to the forecast and the time was within 30 min either side of the forecast period. Table 13 displays the results of the “YES/NO” icing forecasts during two different studies for the BFM and MM5.

Table 13. “YES/NO” icing statistics using BFM and MM5 for all levels and all forecast hours.

Icing Statistics	BFM Study (1999)	MM5 Study (2001)
Samples	112	148
POD	0.66	0.84
FAR	0.13	0.26
Non-event	0.59	0.32
CSI	0.61	0.64
TSS	0.27	0.16
Bias	0.76	1.15

In both models, the same icing routine was used, the original Air Force icing tool with ARL-developed modifications. The results show the same trend from both models; a high POD, a low FAR, and lower TSS due to poor forecast of the non-event; the case of icing being forecasted and no icing having occurred. The most intriguing result is that the BFM has a bias to underforecast the icing cases, while the MM5 tends to overforecast icing. The most likely reason for this is the tendency for the BFM to underforecast moisture and cloud layers, while the MM5 has a moist bias which consequently leads to overforecasting the wintertime cloud layers

and resulting icing. The icing statistics for the BFM with respect to different atmospheric layers are displayed in table 14, while MM5 statistics are shown in table 15.

Table 14. “Yes/No” icing statistics with height for the BFM.

Icing by Height	<=5000 ft AGL	5000–10000 ft AGL	10000–15000 ft AGL
Samples	32	49	26
POD	0.76	0.64	0.65
FAR	0.09	0.10	0.17
Non-event	0.71	0.70	0.00
CSI	0.70	0.59	0.58
TSS	0.47	0.34	–0.35
Bias	0.84	0.72	0.79

Table 15. “Yes/No” icing statistics by height for the MM5.

Icing by Height	<=5000 ft AGL	5000–10000 ft AGL	10000–15000 ft AGL
Samples	37	59	29
POD	0.79	0.86	0.83
FAR	0.32	0.33	0.09
Non-event	0.30	0.27	0.60
CSI	0.58	0.60	0.77
TSS	0.09	-0.14	0.43
Bias	1.16	1.30	0.92

The results in tables 14 and 15 indicate equivalent trends as seen in table 13, where the MM5 is forecasting excessive icing and the BFM underestimating the icing events. While these errors are not extreme, they do cause low TSS values, which can be deceptive in these data since the number of “no” cases is small in almost all the levels studied in this test. More importantly, the POD is high and the FAR low in the BFM reports, with slightly higher POD and higher FAR in the MM5 data. Again, the higher POD in the MM5 is due to the model bias of nearly always forecasting the icing in these layers. The BFM data shows a low FAR which is related to the bias of underforecasting the icing.

While sample size is limited in this icing study, the BFM does have the highest skill and POD in the lowest layers (below 5000 ft AGL), while the MM5 seems to have the most success in the layer from 10000 to 15000 ft AGL. This would agree with the model study, which indicates that the high moisture bias in the MM5 is in the lower levels and that the largest deficit of moisture in the BFM is between 10000 to 15000 ft AGL. In figure 6, the POD of icing is displayed as a function of model forecast time.

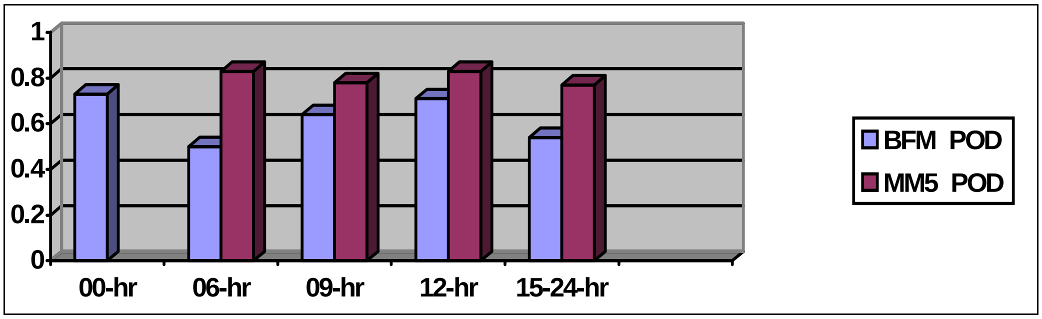

Figure 6. Model icing forecasts by time from the initial model forecast period.

As seen in figure 6, the MM5 has a higher POD from 6 to 24-h due to its bias to overforecast the icing events, while the drier BFM misses more cases of clouds and the resulting icing.

While not shown in a graph or table, the icing routine makes more accurate icing intensity forecasts using the MM5 output than with BFM. The most alarming difference is in the forecasting of moderate icing, where the MM5 had only a 14-precent miss rate compared the BFM's 46 percent miss rate. A careful investigation of the code indicates that this occurs in the case where the temperature of the layer is between –8 to –16 ºC, the dew point depression is less than 1.5 ºC and the lapse rate is greater than 2 ºC/1000 ft. This is the only case where moderate rime can occur in the software. The MM5, with higher moisture values and greater mid-level relative humidity is able to reach this case on more occasions, while the BFM with a dry bias does not represent the mid-level moisture as well and drops into the light icing cases. Another possible problem is in winter when the atmosphere is often colder than –16 ºC, and even if the theory of the icing routine suggests that only light icing can form at such cold temperatures, pilots have reported moderate icing at these temperatures.

In the icing software package, the most significant errors occur in icing types occur in the case where air temperature is between –8 to –16 ºC. According to the decision tree, this is where mixed icing would most likely occur, in cases of unstable lapse rates. However, in the test conducted here, this may be an erroneous assumption about icing formation and occurrence. As already mentioned, most of the icing cases studied transpired when dynamic weather systems caused large-scale lifting of a warmer layer over a colder low level. This creates a favorable environment for stable lapse rates and not unstable lapse rates that might be more common in convective environments. Granted, in some weather situations wintertime precipitation does include convective elements, but in most of the stable winter weather events, the icing routine being used in this study should include the influences of large-scale lifting and stable lapse rates. Some adjustment in the software should account for these conclusions. While this routine is basic and is a proper one for a model without any microphysics package, there must be some adjustments in the software to better determine the icing intensity and icing types.

4.3.3 Cloud Verification

To evaluate the cloud amounts and heights, the cloud forecasts were compared to Meteorological Aviation Routine Weather Reports, which are coded weather observations at selected airports across the world. Since many stations now use automated machines that do not report clouds above 12000 ft AGL, satellite photos were examined to account for higher clouds.

For a forecast to be "correct," the height of the observed cloud had to be within:

- 1000 ft of the forecasted cloud height below 5000 ft AGL
- 1500 ft between 5000 to 10000 ft AGL
- 2000 ft above 10000 ft AGL

Since the error was not considered significant, scattered clouds were not considered wrong forecasts when there was no ceiling forecasted. However, if a ceiling was forecasted and only scattered clouds were observed, the forecast was considered wrong. When a broken layer was forecasted and overcast layer was observed, the forecast was still correct, as was a forecast for overcast conditions where broken clouds were reported. Once an overcast layer was reported, it was impossible to verify any layers above that layer.

For the BFM and MM5 output, the clouds were verified only at the hour of the observation. Table 16 shows the BFM and MM5 cloud forecasts from model initialization period. All time periods for the BFM contained more than 30 samples, while all time periods for the MM5 test included approximately 100 samples.

Table 16. Cloud statistics for the BFM and MM5 by hours from model initialization time.

Hour	BFM (summer 1998) (%)	BFM (winter 1999) (%)	MM5 (winter 2001) (%)
00	68	76	
03	54	61	
06	59	53	56
09	59	50	60
12	65	64	60
18	62	51	55
24	63	49	59
36			46
48			49

NOTE: Values are in percent of correct forecasts.

Results in table 16 indicate that the cloud forecasts follow the pattern that might be expected, with the BFM recording its highest skill in the initial hour and then having the forecast skill decrease with time. The MM5 skill, which was tested in winter only, is nearly identical through

the first 24 h of the model run, although there is decreasing skill in the later time frames. Table 17 shows the cloud statistics by height for the BFM and MM5. The 1999 BFM data are a combination of 0000 and 12000 UTC forecasts, while these 2001 MM5 data are from 0600 UTC model runs.

Table 17. Cloud statistics by height for all model-run hours.

Heights by Hour, Ceilings	<4000 Number of Samples	<4000 Correct (%)	4000 to 8000 Number of Samples	4000 to 8000 Correct (%)	8000 to 20000 Number of Samples	8000 to 20000 Correct (%)
BFM	119	71	40	53	25	27
MM5	370	60	65	10	72	19

Studying these data, the models' post-processing routine has the highest skill in the lowest 4000 ft of the atmosphere, with the BFM correctly handling cloud forecasts 71 percent of the time and the MM5 60 percent of the time. These values decrease with height as both models show limited skill above the boundary layer. While these errors may seem extreme, the cause of the low skill above 4000 ft is mainly due to the both models forecasting cloud layers below these higher levels, thus a ceiling does exist but the forecasts for these ceilings are too low. This trend is more pronounced using the MM5 data, where 63 percent of the "wrong" forecasts were because the cloud routine forecasted a ceiling lower than that observed. In table 18 a listing of ceiling-forecast errors are displayed for the low-cloud observations only. The table indicates the percentages of each error type—forecasts of clouds too low, forecast of clouds too high, and forecasts of no ceiling when one occurred. Tables 19 and 20 show similar statistics but for higher-level ceilings.

Table 18. Model cloud forecast errors for observed ceilings less than 4000 ft AGL.

<4000 ft Cloud Errors	Ceiling Forecast too Low (%)	Ceiling Forecast too High (%)	Ceiling Layer Missed (%)
BFM	26	13	60
MM5	41	46	12

Table 19. Model cloud forecast errors for observed ceilings 4000 to 8000 ft AGL.

4000-8000 ft Cloud Errors	Ceiling Forecasts too Low (%)	Ceiling Forecasts too High (%)	Ceiling Layer Missed (%)
BFM	24	13	62
MM5	77	9	14

Table 20. Mode cloud forecast errors for observed ceilings 8000 to 20000 ft AGL.

8000 to 20000 ft Cloud Errors	Ceiling Forecasts too Low (%)	Ceiling Forecasts too High (%)	Ceiling Layer Missed (%)
BFM	13	13	75
MM5	66	6	28

In the tables 18 to 20, the results show that the BFM and MM5 errors are different but follow some the moisture biases presented in section 4.2. In most of the cases, the BFM error is due to missing the cloud layer, a problem that increases with heights. The MM5 errors are much different, with most of the error due to forecasting the ceilings too low compared to the observed cloud heights. In the lowest levels, less than 4000 ft AGL, the error is more generalized with forecasts divided between ceilings too low and ceilings too high. With increasing height, the MM5 cloud error is most often associated with the ceiling forecast being lower than what was observed.

4.3.4 Visibility and Fog

To evaluate the surface visibilities, the post-processed forecasts were compared to the Meteorological Aviation Routine Weather reports and the Automated Surface Observing System reports across the United States. The entire study was conducted from December 2000 to April 2001 to capture wintertime visibilities

There are seven different forecast classes or ranges for surface visibility, with an emphasis on lower-visibility classes. These ranges are listed below:

1. Less than one mile
2. 1 to 2.99 miles
3. 3 to 4.99 miles
4. 5 to 6.99 miles
5. 7 to 9.99 miles
6. 10 to 19.99 miles
7. 20 miles or greater

All BFM output were from 1200 UTC model runs while all MM5 model runs were from 0600 UTC. There is a built-in bias in this study since many of the model runs were completed on days with forecasted low visibilities and significant storms systems, thus there are more cases of low visibility than what might be expected in a more standardized test.

Surface observations of visibility are very subjective, based on the experience and judgment of the observer or the reliability and accuracy of an instrument. Table 21 displays the “correct” visibility forecasts for each category of both the BFM and MM5.

Table 21. Visibility forecasts, all model-output hours, percent correct for BFM and MM5.

Visibility Observed Values in Miles	BFM (%) Correct	MM5 (%) Correct
<1	41	30
1-2.99	76	44
3-4.99	49	49
5-6.99	40	34
7-9.99	63	80
10-19.99	78	74
>=20	80	88
Total sample (% correct)	59	69

Both models forecast visibility correctly at the higher ranges, with less skill at the lower ranges. Since the most significant aviation problems occur when the visibility is less than 3 miles, these cases are discussed in more detail.

In the low visibility cases, almost all errors are due to the equations missing the low visibility and forecasting the visibility too high. The cause of these errors can be understood by looking at the equations used to forecast the visibility. Looking back at eq 5; the case where a ceiling is known but there is not any precipitation falling, the most significant terms in the regression equation are the surface relative humidity and the surface wind speed. The higher values of relative humidity act to lower the surface visibility, while the stronger wind speeds act to increase the surface visibility. A combination of a high relative humidity along with light surface wind speed would combine to derive the lowest visibility values. This makes sense physically with what is typically observed, however the equations are more sensitive to the relative humidity values at the surface

Forecasting fog is a challenge for mesoscale models since the formation of fog is also dependent upon such elements as droplet size and concentration. Since there is no microphysics package in the BFM, a simple method of just including all forecasted visibilities less than 7 miles as cases of fog. While this assumption is elementary, it is used for both the MM5 and BFM models in this study. Only surface observations with “fog” listed are considered to verify against the forecast of fog. In the results of the “YES/NO” forecasts of fog for all model hours up to 24 h are shown. Beyond 24 h, the MM5 had very low skill scores to detect fog with a POD of 0.47 for 36-h forecasts and only 0.24 for the 48-h forecasts. Much of the error was due to missing the clouds and precipitation events that act to lower the visibility and form fog.

Table 22. Forecast results for fog forecasts to 24 h for the BFM and MM5.

Statistical data	BFM	MM5
Samples	614	399
POD	0.81	0.63
FAR	0.39	0.30
Non-event	0.63	0.84
CSI	0.53	0.50

TSS	0.44	0.47
Bias	1.31	0.89

The trend by model hour is shown in figure 7 which displays the POD (or "YES/NO") forecast of fog for both the BFM and MM5. The BFM data indicate that the initial time period; the 00-hr BFM has the lowest detection of fog. From the 3- to 18-h time period the fog forecasts are all over 80 percent correct. The MM5 trend is dissimilar, with the highest detection of fog in the early model periods with a gradual decrease in skill through the model run.

It can be seen that the BFM initial time has the lowest fog detection (55 percent) compared to the other hours (81 percent for entire sample). The reason for this is probably due to the BFM not forecasting precipitation very well at the initial time period. This will be explained in more detail in the next section, but the problem is related to the model's lower "YES/NO" skill in predicting rainfall at the initial time period. Without the precipitation being correctly forecasted, the visibility forecasts are often higher than observed and the formation of fog is missed.

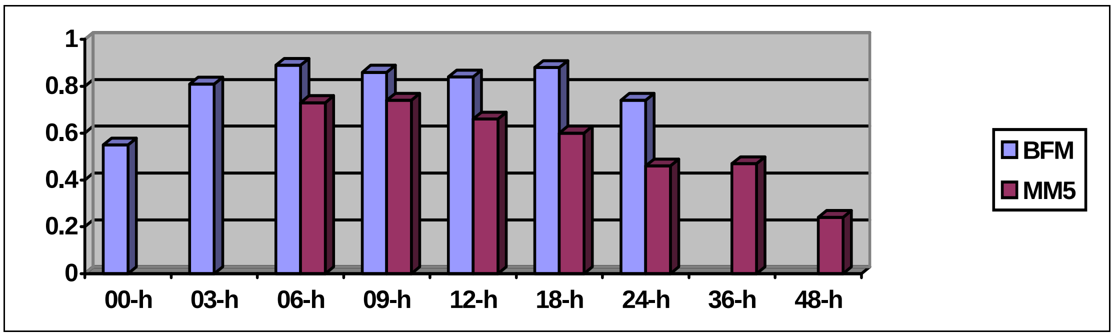

Figure 7. "YES/NO" forecast for fog for both the BFM and MM5, from 0 to 48 h from model initial time.

5. Summary and Discussion

The influence of weather hazards on tactical operations is of great concern to military leaders. These hazards include icing, turbulence, cloud layers, surface visibility, and fog. With the development of an operational mesoscale model, the BFM, short-term weather forecasts (24-h) of model parameters and post-processed are available. Additionally, output from the MM5 to 48-h has furnished ARL with a longer-term operational forecast capacity.

Turbulence is analyzed and forecasted in the ASP by using the PI below 4000 ft AGL, and the TI above 4000 ft AGL. For icing, the RAOB tool originated at AFWA has been modified and is now used in the ARL post-processing. Cloud forecasts were developed through careful investigation of moisture properties on model skew-T diagrams through many different weather environments. This part of the post-processing is the most "rule-based" in its design and uses a series of IF-THEN rules based on relative humidity, height of level, time of the day, season, and

location. The visibility forecasts use a statistically derived regression equation that is utilized in any general weather situation.

This study and report focuses on model evaluation and the influence of the model biases on the post-processed variables. While many studies of model evaluation have been completed and appear in publications, the statistical evaluation of weather hazards, aviation hazards, or post-processed values from mesoscale model output are far less common. Some examples of these evaluations include work done by Marroquin, who discusses statistical verification of aviation–impact variables such as turbulence. In Marroquin's work, he uses the diagnostic turbulence kinetic energy equation using the output from the ETA 30-km, 31-level model during the Storm-scale Operational and Research Meteorology-Fronts Experimental System (STORM-FEST). His results showed a high POD below 5000 ft. AGL but much lower skill above that level. Additionally, Cairns and Chen in STORM-FEST used the Mesoscale Analysis and Prediction System 60-km, 25-level model to analyze "YES/NO" cloud forecasts. They calculated a POD of 0.84 for clouds at heights less than 6500 ft AGL; however, these results include all clouds, unlike the BFM-MM5 study here which are for ceilings only. Brown investigated icing predictions during the Winter Icing and Storms Program in 1994 and found POD values over 0.80, but all FAR over 0.74 for several different routines using the 40-km ETA in levels less than 6000 ft AGL *(32-34)*.

Other aviation-variable studies include one by Dallavalle and Dagostaro which evaluated ceiling and visibility output. Erickson investigated National Weather Service precipitation type routines, while Kim inspected fog forecasting in Korea *(35-37)*.

In the ARL study, there are many examples that show how the derived variables follow the trends of the models. The icing output depends on the dew-point depression from the model-produced vertical layers. Looking at the error in the dew-point depressions of each model, the BFM has an AD of 7.40 ºC at 12 h while the MM5 has an AD of 5.79 ºC at 18 hrs. Since the BFM underforecasts the moisture in 70 percent of the cases, there is a dry bias in the mid-level output of the model. However, the MM5 does not have this dry bias. The BFM has an icing POD of about 0.60 while the MM5 has a POD of 0.83 in the layer where 700 (mb) usually occurs. The icing and clouds are highly dependent on the model moisture forecasts and are strongly correlated since the icing routine will not forecast any icing unless clouds are first forecasted.

One of the most challenging aspects of this project was to create feedback between the derived aviation variables. It is not logical to have low clouds and precipitation with a visibility of 20 statute miles being in the output. There are several cases in the software where checks are made to ensure that there are no inconsistencies of this nature.

As an example of this feedback procedure, in the moist environment, the BFM overforecasts the temperature and underforecasts the dew point at the surface; this is shown in table 5 and the discussion that follows the table. This can lead to a situation where a surface relative humidity is

greatly underforecasted. The MM5 tends to underforecast the temperature and has little bias in the surface dew point, leading to higher relative humidity forecasts than what is observed. This corresponds to the visibility regression equation which indicates that surface relative humidity is the most vital parameter. As a result, it can be expected that the MM5 might forecast lower visibilities more than the BFM. However, this does not occur as seen in table 23.

Table 23. Errors in visibility for the BFM and MM5 in statute miles (sm).

Model/Obstruction	**Fcst Ave (sm)**	**Obs Ave (sm)**	**Mean Absolute Difference (AD)**	**Samples**
BFM No Precip	7.89	9.59	1.98	100
MM5 No Precip	8.35	9.54	1.63	100
BFM fog	5.45	3.30	3.00	68
MM5 fog	5.83	3.36	3.57	79
BFM rain	4.38	3.56	3.21	68
MM5 rain	5.69	3.80	3.39	89
BFM snow	4.71	3.10	3.09	86
MM5 snow	8.12	3.08	5.15	79

In all four cases, the forecast of visibility is higher in the MM5, although in the no-precipitation case the MM5 does have a lower mean absolute difference and a higher percent of correct forecasts. In the fog and rain samples, there is not a significant difference in AD; however, this is not the case when snow was observed to be falling. The most perplexing statistic in table 23 is the high bias in the MM5 snow forecasts. While it has not been verified, the error may be due to a bias in the precipitation rate from the MM5; thus, a low precipitation rate would result in lighter precipitation and less obstruction to the surface visibility in the post-processed feedback checks.

The most unique variable in the post-processing set is turbulence. In eq 2, the main term in the PI is the wind speed while the Turbulence Index (TI) uses the vertical wind shear, convergence, and deformation to calculate clear-air turbulence. Results in the model wind evaluation show little difference or error in the wind speeds; none of the results are significant enough to favor one model over another. Still, the MM5 does have a higher POD, which may be a result of the higher vertical resolution. The icing and turbulence routines seem most influenced by the vertical resolution differences in the models.

The number of interactions of these variables is extremely complicated; however, the most fascinating part of the results is how the output of the post-processing are related to the mesoscale model itself. An error, such as the dry bias in the BFM, can greatly influence the cloud routines and the visibility forecast. This dry bias appears to be related to the failure of the BFM to properly saturate the layers below the cloud layer when precipitation is falling. Other

biases in the radiation parameters, the surface layer, or the model microphysics can be felt in the post-processing.

This study shows that the ARL-derived post-processing routines do work with the AFWA MM5 output, but it is probably best to develop post-processing software for each individual model, which has its own unique trends and biases. The MM5, with higher vertical resolution, performs better for the three-dimensional variables such as icing and turbulence; however, even with only 11 levels above the boundary level the BFM does provide enough information to have positive skill in the higher levels of the model. There appears to be no important influence from the MM5 being a non-hydrostatic model and the BFM using hydrostatic assumptions. Additionally, the differences between a 10- and 15-km grid resolution do not appear to significantly affect any of the model post-processing. It is uncertain if a cloud-microphysics package would have much influence on the aviation variables; however, it may play a role in precipitation types and temperature of the model layers. Since much of this work was conducted during the cold season, the convective parameterization scheme in the MM5 probably plays no role in this study.

In future versions of the BFM, the model will be initialized with MM5 output; thus, it will be intriguing to see if the moist bias of the MM5 has a significant influence when merged with the dry bias of the BFM. Additionally, the BFM will run only to 12-h, with the MM5 being used beyond that time frame. With each change in the model physics or parameterizations the post-processed variables must be researched and studied. While this is one of the disadvantages of using a post-processing technique, it also provides the user with improved results of the derived variables. With improvement in the moisture fields, the models will continue to provide even better skills for many of the variables discussed in this report.

References

1. Passner, J.E. Expert Systems and Empirical Rules for Army Operations on IMETS. Proceedings of the 13th Conference on Weather Analysis and Forecasting, Vienna, VA, 1993, 608-611. (UNCLASSIFIED)

2. Dudhia, J. A Non-hydrostatic Version of the Penn State/NCAR Mesoscale Model: Validation Tests and Simulation of an Atlantic Cyclone and Cold Front, *Monthly Weather Review* **1993**, *121*, 1493-1513.

3. Sauter, D. P. The Integrated Weather Effects Decision Aid (IWEDA): Status and Future Plans, Proceedings of the 1996 Battlespace Atmospheric Conference, Naval Command , Control, and Ocean Surveillance Center, San Diego, CA, 1996, 75-81. (UNCLASSIFIED)

4. Pielke, R.A. Mesoscale Meteorological Modeling, Academic Press, New York, 1984. (UNCLASSIFIED).

5. Yamada, T.; Bunker, S. A Numerical Model Study of Nocturnal Drainage Flows with Strong Wind and Temperature Gradients. *Journal of Applied Meteorology*, **1989,** *28*, 545-554.

6. Henmi, T.; Dumais R.E. Jr., *Description of the Battlescale Forecast Model*, ARL-TR-569, U.S. Army Research Laboratory, White Sands Missile Range, NM, 1998. (UNCLASSIFIED)

7. Passner, J.E.; Henmi, T. Precipitation Forecasts Using the Battlefield Forecast Model, Symposium on Precipitation Extremes: Prediction, Impacts, and Responses, Albuquerque, NM, 2001395-397. (UNCLASSIFIED)

8. Sasamori, T. The Radiative Cooling Calculation for Application to General Circulation Experiments. *Journal of Applied Meteorology*, **1968**, *7*, 721-729.

9. Sundqvist, H.; Berge, E.; Kristja'nsson, J.E. Condensation and Cloud Parameterization Studies with a Mesoscale Numerical Weather Prediction Model, *Monthly Weather Review*, **1989,** *117* 1641-157.

10. Grell, G.A.; Dudhia, J.; Stauffer, D.R. *A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5)*, NCAR Tech. Note NCAR/TN-398+STR; 1995. (UNCLASSIFIED)

11. Bresh, J.; et al. On the Use of the MM5 in an Aviation Weather Forecast System. Ninth Conference on Aviation, Range, and Aerospace Meteorology, Orlando, Fl, **2000,** 167-170. (UNCLASSIFIED)

12. *Glossary of Meteorology*, Fifth Edition, American Meteorological Society, Boston, MA, 1959.

13. Rammer, W.A. Model Relationship of CAT to Upper Wind Flow Patten, National Meteorological Center Aviation Weather Forecast Branch, Note 14, 1973. (UNCLASSIFIED)

14. Lake, H. Analysis of Clear Air Turbulence. *Geophys. Res. Rep*., No 47, Air Force Cambridge Research Center, 63, **1956**.

15. Black, T.L.; Marroquin, A. Preliminary Results of Turbulence Predictions for use in Aviation Weather Forecasts. Preprint of The Fifth International Aviation Weather Systems, Vienna, VA, **1993;** 461-462.

16. Miles, J.W.; Howard, L.N. Notes on Heterogenous Shear Flow. *J. Fluid Mech*, **1964,** 20, 331-336.

17. Stull, R. B. *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers: Boston, MA, 1989.

18. Boyle, J.S. *Turbuelnce Indices Derived From FNOC Field and TOVS Retrievals*; NOARL Technical Note 47; Naval Oceanographic and Atmospheric Research Laboratories: 1990. (UNCLASSIFIED)

19. Elrod, G.P.; Knapp, D.I. An Objective Clear-Air Turbulence Forecasting Technique: Verification and Operational Use. *Wea. Forecasting* **1992**, 7, 150-165.

20. Petterssen, S., *Weather Analysis and Forecasting Vol. 1*, McGraw-Hill, 1956.

21. Knapp, D. I.; Smith, T. J.; Dumais, R. Development and Verification of a Low-Level Turbulence Analysis and Forecasting Index Derived From Mesoscale Model Data. Preprint of *Sixth Conference on Aviation Weather Systems*, Dallas, TX, 1995, 436-440. (UNCLASSIFIED)

22. Passner, J.E. *An Evaluation of the Three-Dimensional Weather Hazards Using Sounding Data and Model Output Data*; ARL-TR-1046; U.S. Army Research Laboratory; White Sands Missile Range, NM, 2000. (UNCLASSIFIED)

23. *Forecasters Guide on Aircraft Icing;* Air Weather Service, AWS/TR-80/001; Air Weather Service, Scott Air Force Base, IL, 1980, 61.

24. Knapp, D.I., *Verification Report: Comparison of Various Icing Alaysis and Forecasting Techniques;* Air Force Global Weather Central: Offutt Air Force Base, NE, 1992. Unpublished work.(UNCLASSIFIED)

25. Cornell, D.; Donahue, C.A.; Chan, K. *A Comparison of Aircraft Icing Forecast Models*, AFCCC/TN-95/004, U.S. Air Force Combat Climatological Center: Scott Air Force Base, Il, 1995, 40. (UNCLASSIFIED)

26. Walcek, C. J. Factors Influencing Regional-scale Cloud Cover: Investigations Using Satellite-Derived Cloud Cover and Standard Meteorological Observations. Proceedings of the Fourth Symposium of Global Change Studies, Anaheim, CA 1993, 235-236. (UNCLASSIFIED)

27. Schultz, P.; Polotovich, M.K. Toward the Improvement of Aircraft-Icing Forecasts for the Continental United States. *Weather and Forecasting* **1992**, 7, 491-500. (UNCLASSIFIED)

28. Knapp, D.I. Development of a Surface Visibility Algorithm for Worldwide Use with Mesoscale Model Output. 15th Conference of Weather Analysis and Forecasting, Norfolk, VA,. 1996, 83-86. (UNCLASSIFIED)

29. Donaldson, R.J.; Dyer, R.M.; Krauss, M.J. An Objective Evaluator of Techniques for Predicting Severe Weather Events, Preprints of *Ninth Conference Severe Local Storms*, Norman, OK, 321-326, 1975. (UNCLASSIFIED)

30. Woodcock, F. The Evaluation of Yes/No Forecasts for Scientific and Administrative Purposes. *Mon. Wea. Rev* **1976,** *104*, 1209-1214.

31. Kane, T.; Brown, B.G.; Sharman, R. A Preliminary Climatology of Upper Level Turbulence Reports, Preprints of *8thConference on Aviation, Range, and Aerospace Meteorology*, Dallas, TX, 1999363-367. (UNCLASSIFIED)

32. Marroquin; et al. Importance of Statistical Verification of AIVs From Mesoscale Numerical Model Output. Preprint of *6thConference of Aviation Weather Systems*, Dallas, TX, 1995198-203. (UNCLASSIFIED)

33. Cairns, M. M.; Chen, J. A Second Evaluation of Aviation-Impact Variables Generated by Numerical Systems. Preprint of *6thConference of Aviation Weather Systems*, Dallas, TX, 1995, 204-206. (UNCLASSIFIED)

34. Brown; et al. WISP94 Real-Time Icing Prediction and Evaluation Program (WRIPEP): Statistical Issues and Forecast Verification Results. Preprint of *6thConference of Aviation Weather Systems*, Dallas, TX, 1995, 207-212, (UNCLASSIFIED)

35. Dallavolic, J.; Dagostaro, P.; Dagostaro, V. Accuracy of Ceiling and Visibility Forecasts Produced by the National Weather Service. Preprint of *6thConference of Aviation Weather Systems*, Dallas, TX, 1995, 213-218. (UNCLASSIFIED)

36. Erickson, M. C. Evaluating NWS Precipitation Type Forecasts. Preprint of *6thConference of Aviation Weather Systems*, Dallas, TX, 1995, 219-224. (UNCLASSIFIED)

37. Kim, J.; et al. Fog Forecast for Kimpo International Airport of Korea. Preprint of 9th *Conference of Aviation Weather Systems*, Orlando, FL, 2000, 219-222. (UNCLASSIFIED)

Acronyms and Abbreviations

AD	absolute difference
ADI	Alternating Direction Implicit
AFWA	Air Force Weather Agency
AFGWC	Air Force Global Weather Center
AGL	above ground level
AI	artificial intelligence
ARL	Army Research Laboratory
ASP	Atmospheric Sounding Program
BFM	Battlescale Forecast Model
CC	correlation coefficient
CAT	clear-air turbulence
CIG	ceiling
CNE	correct non-event
CONUS	continental United States
CSI	critical success index
FAR	false alarm rate
FNMOC	The U.S. Navy Fleet Numerical Meteorological and Oceanography Center
GMDB	gridded meteorological database
GriB	gridded binary form
HOTMAC	Higher Order Turbulence Model for Atmospheric Circulations
IMETS	Integrated Meteorological System
IWEDA	Integrated Weather Effects Decision Aids
mb mbar	millibar
MM5	Pennsylvania State University/National Center for Atmospheric Research Mesoscale Model Version 5

MMPOST	MM5 Post-processing output
TSS	true skill score
3-D	three dimensional
NCAR	National Center for Atmospheric Research
NOGAPS	Naval Operational Global Atmospheric Prediction System
PI	Panofsky index
PIREP	pilot reports
POD	probability of detection
RAOB	radiosonde upper-air observation
RI	Richardson Number
RMSE	root-mean square error
sm	statute miles
STORM-FEST	Storm-scale Operational and Research Meteorology -Fronts Experimental System
TD	dew point temperature
TDA	Tactical Decision Aids
UTC	universal time coordinates
VISCAT	visibility category